Vol. 123, No. 1, 1984 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 30, 1984 Pages 1-7

HUMAN APOLIPOPROTEIN A-II: NUCLEOTIDE SEQUENCE OF A CLONED cDNA, AND LOCALIZATION OF ITS STRUCTURAL GENE ON HUMAN CHROMOSOME 1

Marsha N. Moore[1], Fa-Ten Kao[2], Yuan-Kai Tsao[1], and Lawrence Chan[1]

[1]Departments of Cell Biology and Medicine,
Baylor College of Medicine, Houston, TX 77030

[2]Eleanor Roosevelt Institute for Cancer Research
and Department of Biochemistry, Biophysics, and Genetics,
University of Colorado, Health Sciences Center,
Denver, Colorado 80262

Received July 12, 1984

SUMMARY: Using the technique of oligonucleotide hybridization, we have isolated two dscDNA clones to human apolipoprotein A-II. One of the clones (pAII-1) has been completely sequenced. It has 433 nucleotides which includes a poly A tail of 10 adenylic acid residues, all the coding and 3'-non-translated regions of the mRNA and part of the 5'-non-translated region. The amino acid sequence derived from the cDNA clone includes 100 amino acids (including the 23 amino acid prepropeptide) which is very similar to the sequence reported by Brewer et al. (Proc. Natl. Acad. Sci. USA 69, 1304-1308). [^{32}P]-labeled pAII-1 was used as a probe in chromosome mapping studies to detect the human apoAII structural gene sequence in human-Chinese hamster cell hybrids. Southern blot analysis of 10 hybrids localized the gene to human chromosome 1.

Plasma high density lipoproteins (HDL) represent a spectrum of lipoprotein particles with a density between 1.063 and 1.250 g/ml. Two apolipoproteins, apoA-I and apoA-II, comprise about 90% of the total protein mass in HDL. ApoA-I, the most abundant protein in HDL, consists of 243 amino acids with a molecular weight of approximately 28,000 (1,2). ApoA-II is a dimer of identical polypeptides each 77 residues in length joined by a disulfide bond on residue 6 (3). It is present at about 20-25% the concentration of apoA-I in human HDL. Although the functional role of apoA-II remains uncertain, there is some evidence that it might regulate the activities of lecithin-cholesterol acyltransferase (4) and hepatic lipase (5,6).

The biosynthesis of human apoA-II has been studied by Stoffel et al. (7) and by Gordon et al. (8). These studies indicate that the protein is initially synthesized as a preproprotein. The lengths of the signal peptide and prosegment, however, are different in the two reports. In this communication, we

Abbreviations: HDL, high density lipoproteins; kb, kilobase(s); apo, apolipoprotein.

0006-291X/84 $1.50

report the cloning and complete sequence analysis of human apoA-II cDNA. Furthermore, by the method of Southern blot analysis of human-Chinese hamster somatic cell hybrids, we have localized the apoA-II structural gene to human chromosome 1.

EXPERIMENTAL PROCEDURES

cDNA Library

A human cDNA library was a generous gift of Drs. S.L.C. Woo and T. Chandra. It was constructed from a mRNA preparation from adult liver tissue. The ds cDNA was inserted into the Pst I site of pBR322 by the GC-tail technique. The host was Escherichia coli K12 RRI.

Oligonucleotide Synthesis and Labeling

A mixture of oligonucleotides (17-mers) with the sequences shown in Figure 1 was custom synthesized by P-L Biochemicals by the phosphotriester technique (9). The sequences of the oligonucleotides were confirmed by nucleotide sequencing using the Maxam-Gilbert technique (10). Oligonucleotides were labeled at the 5' end by transfer of ^{32}P from [γ-^{32}P]ATP by using bacteriophage T4 polynucleotide kinase (Boehringer-Mannheim) as described (10). Labeled oligonucleotides were purified by DE52 ion exchange chromatography (11).

Screening of cDNA Clones

Recombinant cDNA clones were replica plated on tetracycline (25 µg/ml) plates, transferred to Whatman 541 filter paper, amplified with chloramphenicol, and prepared for hybridization as described by Suggs *et al.* (12) and Wallace *et al.* (11). Temperature of hybridization was 37°C.

Restriction Mapping and Nucleotide Sequence Determination

Inserted cDNA sequences were recovered by digestion with Pst 1. Restriction mapping was performed by standard techniques under conditions recommended by the suppliers. The entire sequence of pAII-1 was determined by a combination of the chemical method of Maxam and Gilbert (10) and the dideoxynucleotide chain termination method of Sanger *et al.* (13) following subcloning in the M13 vectors mp8 or mp9 (14).

Localization of ApoAII Structural Gene By Somatic Cell Hybrids

Mapping of apoA-II structural gene on human chromosomes was performed by Southern blot analysis of DNAs isolated from a panel of human-Chinese hamster somatic cell hybrids (15,16). The conditions for culture of the parental cell

	12					17	
Amino Acid	Ser	Gln	Tyr	Phe	Gln	Thr	
mRNA	5'(UG/AC C/A A/G	CA A/G	UA U/C	UU U/C	CA A/G	AC U/C/A/G	)3'
Oligonucleotide	3'd(AC/TG G/C/T	GT I/C	AT A/G	AA A/G	GT I/C	TG	)5'

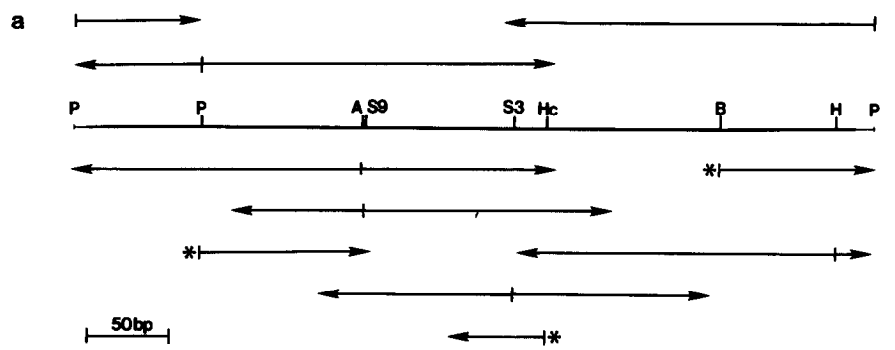
Figure 1: Sequence of oligonucleotides used as hybridization probe. The last wobble for Thr 17 was dropped to reduce the number of oligonucleotides by 4-fold. This portion of the amino acid sequence is based on that published by Brewer *et al.* (3).

lines (CHO-K1, and HT-1080), fusion between the auxotrophic mutants of CHO-K1 cells and human cells, and the characterization of the panel of somatic cell hybrids with respect to their human chromosome content using cytogenetic and isozyme complements have been described previously (15-19). The cell hybrids used in the present study are subpopulations derived from the respective original hybrids. These hybrid subpopulations have been characterized by cytogenetic and isozyme analyses at about the same time the DNAs were prepared. The cytogenetic analysis involved sequential staining of the same chromosome slides with trypsin-banding and Giemsa-11 differential staining procedures (18,19). The human chromosome content in some hybrids has also been confirmed by isozyme analysis. Due to the continued growth in culture for some hybrids, and thawing of the frozen cultures for other hybrids, some changes in the human chromosome content occurred in the subpopulations as compared to the original hybrids. The percentage of each human chromosome retained in the hybrid subpopulations generally ranged from 50-100% of the cells analyzed.

For DNA preparations, cells were grown in 150 mm dishes to confluency and harvested by trypsinization. They were treated with proteinase K after washing and DNA was isolated as described (15,16). Digestion and agarose gel electrophoresis of cell hybrid DNA and Southern blot analysis were performed as described in Cheung et al. (16).

RESULTS AND DISCUSSION

By the method of oligonucleotide hybridization, we have identified 7 positive colonies after screening about 50,000 colonies. Two of these have identical restriction maps and the clone with a longer cDNA insert, pAII-1, was completely sequenced. The sequencing strategy and complete nucleotide sequence of this clone are shown in Figures 2a and 2b. The clone contains a cDNA insert of 433 nucleotides which include the poly A tail, the complete 3'-non-translated region of 115 nucleotides, the complete coding region of 300 nucleotides, and 8 nucleotides of the 5'-non-translated region. The coding portion predicts a peptide of 100 amino acids. The mature plasma apo-AII peptide starts at residue 24, and consists of 77 amino acid residues. This part of the derived amino acid sequence completely matches the amino acid sequence published by Brewer et al. (3), with one exception. Residue 37 of the mature protein is predicted to be Glu rather than Gln as reported previously (3). As this part of the DNA sequence was determined on both strands on four restriction fragments by the chemical degradation method, and also matched the sequence obtained by the M13 dideoxynucleotide chain termination method, we are confident that the codon for this amino acid is GAG. Whether the difference observed represents true sequence heterogeneity of human apoA-II remains to be determined. If it is proven to be the case, it may be of



b

```
1   GTT ACC AAC ATG AAG CTG CTC GCA GCA ACT GTG CTA CTC CTC ACC ATC TGC AGC CTT GAA GGA GCT TTG GTT CGG AGA
          MET LYS LEU LEU ALA ALA THR VAL LEU LEU LEU THR ILE CYS SER LEU GLU GLY ALA LEU VAL ARG ARG

79  CAG GCA AAG GAG CCA TGT GTG GAG AGC CTG GTT TCT CAG TAC TTC CAG ACC GTG ACT GAC TAT GGC AAG GAC CTG ATG
    GLN ALA LYS GLU PRO CYS VAL GLU SER LEU VAL SER GLN TYR PHE GLN THR VAL THR ASP TYR GLY LYS ASP LEU MET

157 GAG AAG GTC AAG AGC CCA GAG CTT CAG GCC GAG GCC AAG TCT TAC TTT GAA AAG TCA AAG GAG CAG CTG ACA CCC CTG
    GLU LYS VAL LYS SER PRO GLU LEU GLN ALA GLU ALA LYS SER TYR PHE GLU LYS SER LYS GLU GLN LEU THR PRO LEU

235 ATC AAG AAG GCT GGA ACG GAA CTG GTT AAC TTC TTG AGC TAT TTC GTG GAA CTT GGA ACA CAG CCT GCC ACC CAG TGA
    ILE LYS LYS ALA GLY THR GLU LEU VAL ASN PHE LEU SER TYR PHE VAL GLU LEU GLY THR GLN PRO ALA THR GLN .

313 AGT GTC CAG ACC ATT GTC TTC CAA CCC CAG CTG GCC TCT AGA ACA CCC ACT GGC CAG TCC TAG AGC TCC TGT CCC TAC

391 CCA CTC TTT GCT ACA ATA AAT GCT GAA TGA ATC CAA AAA AAA AA
```

Figure 2a: Partial restriction map and sequencing strategy of pAII-1. The arrows above the map indicate sequences obtained by the M13 method. The arrows below the map indicate sequences obtained by the Maxam-Gilbert technique. * indicates 5' end labeling. The other fragments were sequenced by 3' end labeling as described in Experimental Procedures. P, Pst 1; A, Ava II; S9, Sau 961; S3, Sau 3A1; Hc, Hinc II; B, Bal 1; H, Hinf 1. b. Nucleotide sequence of pAII-1. The deduced amino acid sequence is indicated. The mature plasma protein starts with first residue on the second line, i.e., Gln. The prepro-segments are represented by the 23 residues on the first line, ending with Arg. The putative polyadenylation signal, AATAAA, is underlined.

functional significance since it changes the charge properties of the apolipoprotein.

The first 23 residues of the derived amino acid sequence correspond to the prepropeptide of human apoA-II. This part of the sequence completely matches the partial sequence reported by Gordon et al. (8).

Using cloned human apo-AII cDNA as a hybridization probe, we have localized the structural gene for this apolipoprotein on human chromosome 1 by the method of Southern blot analysis of DNA from a panel of human-Chinese hamster somatic cell hybrids. Pilot experiments demonstrated that the nick-translated pAII-1 hybridized to an ~ 12 kb band on a Southern blot analysis of Eco R1 digested human DNA (Figure 3). Under the conditions of hybridization, no cross-hybridizing Chinese hamster DNA fragments were detected. Similarly, when pAII-1 was used to probe a Southern blot of Bam H1 digested DNA, an ~ 10

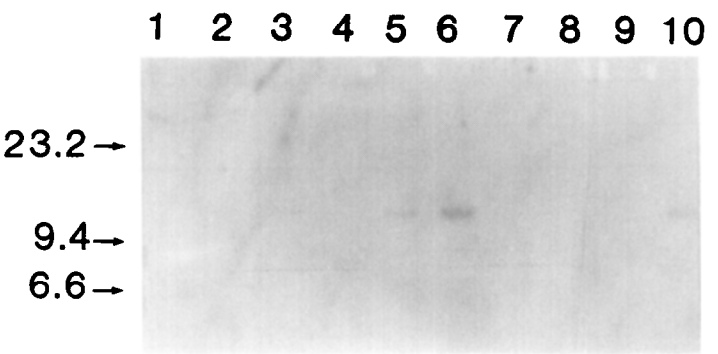


Figure 3: Hybridization of pAII-1 to Eco Rl-digested DNA from human-Chinese hamster somatic cell hybrids and their parental lines. Lanes 1-8, different somatic cell hybrids. Lane 9, Chinese hamster cell CHO-Kl. Lane 10, human cell HT 1080. Lanes 3,5,6 are positive for the human specific 12 kb DNA band. They all containing human chromosome 1. Lanes 1,2,4,7 and 8 are negative, none of which contain human chromosome 1. For complete tabulation of results, please see Table I.

kb band was detected for human DNA, but no cross hybridization was observed for Chinese hamster DNA digested with the same enzyme (data not shown). We have used Southern blots of hybrid cell DNA digested with Eco Rl or Bam Hl and hybridized them to pAII-1 and correlated the presence of the human specific DNA bands with the presence of specific human chromosomes in the various somatic cell hybrids.

Table 1 shows the results of Southern blot analysis with [^{32}P]-labeled pAII-1 hybridized to the DNAs from 10 human CHO-Kl cell hybrids. Synteny analysis of the various hybrid clones indicates concordance between the apoA-II gene and human chromosome 1. Thus, the presence or absence of the human specific hybridizing DNA band correlates only with human chromosome 1. These results allow assignment of the structural gene for human apoA-II to chromosome 1 and to no other chromosomes (Table 1).

To date, the genes for a number of apolipoproteins have been assigned to specific human chromosomes. Both apoE and apoC-II have been assigned to human chromosome 19 (20,21), and apoA-I has been assigned to the long arm of human chromosome 11 (16). Furthermore, the gene for familial hypercholesterolemia (22) and the low density lipoprotein receptor (23) has also been mapped to chromosome 19. We have now mapped the gene for another apolipoprotein, apoA-II, and shown that it is on human chromosome 1. It is interesting that in the

Table I. Synteny Analysis of the Human apoA-II Gene in 10 Human CHO-K1 Cell Hybrids by Molecular Hybridization Using a cDNA Probe pAII-1

Hybrids	Human Chromosomes*																							Apo A-II@
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
CP3-1	-	-	-	+	+	-	-	-	-	-	+	+	-	+	-	+	+	+	-	+	+	-	+	-
CP4-1	-	-	-	+	+	-	-	+	-	-	+	-	-	+	-	-	-	-	-	-	-	+	+	-
CP5-1	+	-	-	-	+	-	-	+	+	-	-	+	-	+	+	-	+	-	+	-	+	+	-	+
CP6-1	-	-	-	+	-	-	-	-	-	-	-	+	-	+	-	-	+	+	+	-	+	+	-	-
CP12-1	-	+	-	+	-	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	+	-
CP14-1	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	+	+	-
CP16-1	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-	+	-	+	+	+	+	-	-
CP20-1	-	+	-	-	+	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-
CP26-1	+	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-	-	+	-	-	+
CP28-1	+	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	+	+	+	-	-	-	+
Number of Concordant Hybrids	10	5	7	4	5	8	8	7	8	7	5	6	7	4	8	6	4	6	7	6	4	3	3	
Percent of Concordance	100	50	70	40	50	80	80	70	80	70	50	60	70	40	80	60	40	60	70	60	40	30	30	

*Identified by trypsin-banding and Giemsa-11 differential staining in sequential steps and by isozyme analysis.

@Identified by Southern blot analysis using an apoA-II cDNA probe, pAII-1.

mouse, apoA-I was mapped to chromosome 9 (24), to a region highly homologous to the long arm of human chromosome 11 (16,25,26), whereas apoA-II was mapped to mouse chromosome 1 (24). Whether segments of human chromosome 1 next to the apoA-II gene are homologous to similar segments on mouse chromosome 1 on which mouse apoA-II DNA resides remains to be determined.

ACKNOWLEDGMENTS

This investigation was supported by grants from the National Institutes of Health, HL-27341 and HD-02080, and the March of Dimes Birth Defects Foundation. We thank Ms. Sundi Chen for excellent technical assistance.

REFERENCES

1. Brewer, H.B., Jr., Fairwell, T., LaRue, A., Ronan, R., Houser, A., and Bronzert, T.J. (1978) Biochem. Biophys. Res. Commun. 80, 623-630.
2. Cheung, P., and Chan, L. (1983) Nucleic Acids Res. 11, 3703-3715.
3. Brewer, H.B., Lux, S.E., Ronan, R., and John, K.M. (1972) Proc. Natl. Acad. Sci. USA 69, 1304-1308.

4. Scanu, A.M., Lagocki, P., and Chung, J. (1980) Ann. N.Y. Acad. Sci. 348, 160-173.
5. Jahn, C.E., Osborne, J.C., Jr., Schaefer, E.J., and Brewer, H.R., Jr. (1981) FEBS Lett. 131, 366-368.
6. Shinomiya, M., Sasaki, N., Barnhart, R.L., Shirai, K., and Jackson, R.L. (1982) Biochim. Biophys. Acta 713, 292-299.
7. Stoffel, W., Kurger, E., and Deutzmann, R. (1983) Hoppe Seyler's Z. Physiol. Chem. 364, 227-237.
8. Gordon, J.I., Budelier, K.A., Sims, H.F., Edelstein, C., Scanu, A.M., and Strauss, A.W. (1983) J. Biol. Chem. 258, 14054-14059.
9. Itakura, K., and Riggs, A.D. (1980) Science 209, 1401-1407.
10. Maxam, A.M., and Gilbert, W. (1980) in Methods in Enzymology, Moldave, K., and Grossman, L., eds., Vol. 65, pp. 499-560, Academic Press, New York.
11. Wallace, R.B., Schold, M., Johnson, M.J., Dembek, P., and Itakura, K. (1981) Nucleic Acids Res. 9, 3647-3656.
12. Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H., and Itakura, K. (1981) Proc. Natl. Acad. Sci. USA 78, 6613-6617.
13. Sanger, F., Coulson, A., Barrell, B., Smith, A., and Roe, B. (1980) J. Mol. Biol. 143, 161-178.
14. Messing, J., and Vieira, J. (1982) Gene 19, 269-276.
15. Kao, F.T., Hartz, J.A., Law, M.L., and Davidson, J.N. (1982) Proc. Natl. Acad. Sci. USA 79, 865-869.
16. Cheung, P., Kao, F.T., Law, M.L., Jones, C., Puck, T.T., and Chan, L. (1984) Proc. Natl. Acad. Sci. USA 81, 508-511.
17. Kao, F.T. (1973) Proc. Natl. Acad. Sci. USA 70, 2893-2898.
18. Kao, F.T., Jones, C., and Puck, T.T. (1976) Proc. Natl. Acad. Sci. USA 73, 193-197.
19. Morse, H.G., Patterson, D., and Jones, C. (1982) Mamm. Chromosome Newsl. 23, 127-133.
20. Olaisen, B., Teisberg, P., and Gedde-Dahl, T., Jr. (1982) Human Genet. 62, 233-236.
21. Jackson, C.L., Bruns, G.A.P., and Breslow, J.L. (1984) Proc. Natl. Acad. Sci. USA 81, 2945-2949.
22. Berg, K., and Heiberg, A. (1978) Cytogenet. Cell Genet. 22, 621.
23. Francke, U., Brown, M.S., and Goldstein, J.L. (1984) Proc. Natl. Acad. Sci. USA 81, 2826-2830.
24. Lusis, A.J., Taylor, B.A., Wangenstein, R.W., and LeBouef, R.C. (1983) J. Biol. Chem. 258, 5071-5078.
25. Meisler, M.H., Wanner, L., Kao, F.-T., and Jones, C. (1981) Cytogenet. Cell Genet. 31, 124-128.
26. Wang, A.L., Arredondo-Vega, F.X., Gaimpietro, P.F., Smith, M., Anderson, W.F., and Desnick, R.J. (1981) Proc. Natl. Acad. Sci. USA 78, 5734-5738.